Route Descriptions

Baltaretu, A.A.; Krahmer, E.J.; Maes, Alfons

Publication date:
2014

Document Version
Peer reviewed version

Route descriptions: the role of intersection type and visual clutter for spatial reference

Adriana Baltaretu, Emiel Krahmer, and Alfons Maes

Tilburg Center for Cognition and Communication (TiCC), Tilburg University
{adriana.baltaretu,e.j.krahmer,maes}@tilburguniversity.edu

Keywords: route directions, visual clutter, intersection type, landmarks

1 Introduction

New technological advances (e.g., Google Glasses) enable context aware pedestrian navigation systems to generate instructions making use of all (variable and stable) environmental information. We know little about how the visual surroundings influence the turn-by-turn production of pedestrian navigation instructions. We address this issue by analysing the effects of environment complexity on reference to landmarks and paths. In this study, complexity was operationalized as the intersection structure and the richness of details in the visual scene. We expect route descriptions (RDs) to contain more environmental information as the visual surrounding becomes more complex.

Route descriptions at least include an action coupled with direction (*go left*) as well as path information (*first street*) and can be enhanced by landmarks (*at the pharmacy*). These instructions are hypothesised to vary depending on the (geometrical) structure of the intersection [2]. In simple intersections street branches are intersecting at 90° angle, the number of turning options is quite limited and the level of uncertainty is low. Thus, we would expect instructions to include a minimum amount of information (e.g., reference to action, direction and path). The complexity of an intersection increases with the number of branches, the intersecting angle, and the options of turning (e.g., turn right in a K - shaped intersection). In such situation, we would expect people to produce detailed descriptions (more path references) and more references to salient entities, such as landmarks, that can make the route description easier to understand.

In addition, we hypothesise that RD production is also influenced by low level visual factors such as the amount of detail in a scene. Language and vision are interrelated in a complex way and detailed (cluttered) visual environments have been shown to influence language production. For example, the more visual details is in a scene, the longer it took respondents to start typing their scene descriptions, and cluttered scenes resulted in the production of more complex constructions [1]. Congruent evidence comes from a memory based route production task where the number of landmarks was kept constant across conditions [3]. In this study, it was found that participants that had previously seen cluttered scenes more often added landmarks in their RDs, than those who had seen maps with lower levels of visual detail. We expect that when scenes are harder

to process, subjects would produce more detailed instructions (more references to landmarks) and longer instructions.

2 Method

2.1 Participants

78 participants were paid to take part in the experiment via a crowdsourcing service. After excluding the non-native English speakers, the final sample that was analysed included 43 participants (15 males, mean age 44 years).

2.2 Materials

A pool of approximately 200 scenes with a pedestrian street based view was created by taking snapshots of rural and urban intersections in Google StreetView. Two scene types were created: simple (T- and +- shaped) and complex intersections (Y- and K– shaped, as well as crossroads with 5 branches). The level of visual clutter in these pictures was estimated using the Feature Congestion algorithm [4] and human ratings. The final set of stimuli consisted of 36 scenes (see examples in Fig.1, Fig.2). Yellow lines depicting the route and the direction to be followed (left, right and straight) were drawn using an open source graphics editor.

2.3 Procedure

The instructions specified the scenario stating that we are developing software that can generate real time/live pedestrian route descriptions based on the visual input coming from the Google Glasses video camera and realized in audio format via a smartphone. The task for participants was to provide route instructions. Participants saw one picture at a time and filled in the description in the input

Fig. 1: Example of simple / complex intersections in scenes with a low level of visual clutter

Fig. 2: Example of a simple / complex intersections in scenes with a high level of visual clutter

field provided under the picture. The task started with 3 warm-up trials, than 36 experimental trials were presented in random order. Lastly, they filled in a series of demographic questions.

3 Results

The RDs (N 36*43 = 1548) were coded for presence of landmarks (references to visual objects), path references (references to channels of movement) and description length. The RD components were analysed separately using logit mixed model analysis with Clutter and Intersection type as fixed factors; participants and item pictures as random factors; p - values were estimated via parametric bootstrapping.

3.1 Landmarks reference

For the number of landmarks there was no main effect of Intersection type, ($p > .05$) . There was a main effect of Clutter ($\beta = .87$, $SE = .32$, $p < .01$). RDs in low cluttered scenes had fewer landmarks ($M = .11$) compared to scenes with high clutter levels ($M = .25$). In addition, there was a significant interaction ($\beta = .62$, $SE = .44$, $p < .05$) between the main factors: low cluttered scenes triggered in both types of intersection similar numbers of references ($M = .11$) (see Fig.3). In high cluttered scenes there are more landmark references in complex intersections ($M = .31$) than in simple intersections ($M = .17$).

3.2 Path reference

For the number of path references there was a main effect of Intersection type ($\beta = .39$, $SE = .16$, $p < .05$). Unsurprisingly, simple intersections trigger less path references ($M = .58$), than complex intersections ($M = .97$), where more

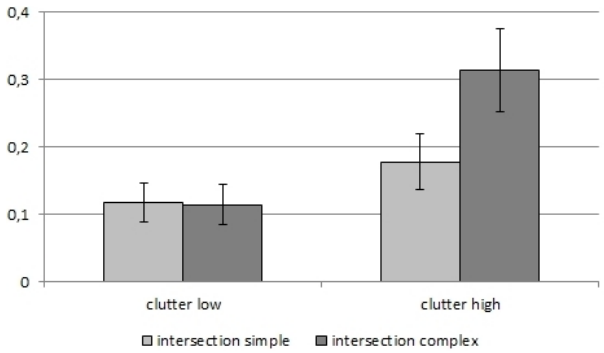

Fig. 3: Average number of landmarks per scene description as a function of Clutter and Intersection type (error bars indicate standard error)

paths are involved. There was no effect of Clutter ($p > .05$) and no interaction between the two factors ($p > .05$).

3.3 Length of descriptions

For the overall number of words in the route descriptions there was a main effect of Intersection type ($\beta = .29$, $SE = .12$, $p < .05$). RDs in simple intersections ($M = 5,15$) are shorter than those in complex intersections ($M = 8,00$). There was no effect of Clutter ($p > .05$) and no interaction between the two factors ($p > .05$).

4 Discussion

In this paper, we have investigated how intersection type and visual clutter influence the length of instructions, landmark and path references. Clutter affected the number of landmark references (high cluttered scenes contained a larger number of references), while intersection type influenced the number of path references and description length. Of interest for context aware system development, these results highlight that not only the complexity of the intersection, but also the overall level of visual clutter in the environment play a role in the production of route directions.

References

1. Coco, M. I., Keller, F. Sentence production in naturalistic scenes with referential ambiguity. In Proc. 32th CogSci Conference, Portland, (2010)
2. Hirtle, S., Richter, K. F., Srinivas, S., Firth, R. This is the tricky part: When directions become difficult. Journal of Spatial Information Science, (1), 53–73, (2014)
3. Westerbeek, H., Maes, A. Routeexternal and Routeinternal Landmarks in Route Descriptions: Effects of Route Length and Map Design. Applied Cognitive Psychology, 27(3), 297–305, (2013)
4. Rosenholtz, R., Li, Y., Nakano, L. Measuring visual clutter. Journal of Vision, 7(2), 17, (2007)